All cats are gray in the dark: enrichment/depletion approaches for biomarker discovery on felis catus plasma

Carcangiu, L ; Pisanu, S ; Tore, S ; Addis, M F ; Zini, Eric ; Uzzau, S ; Pagnozzi, D

Abstract: In veterinary medicine, assay performance is often affected by the lack of species-specific diagnostic tools. Reliable biomarkers might be identified by investigating biological fluids of the species of interest, but protein sequence databases are often incomplete and human-specific devices for reducing sample complexity might fail when applied to animal plasma. Here, seven commercial methods based on different capturing agents (anti-human antibodies, affinity ligands, mixture of antibodies and ligands, and combinatorial peptide ligand libraries) are applied to cat plasma and evaluated in terms of yield, identified proteins/ peptides, and relative abundance by high-resolution shotgun proteomics and label-free quantitation. As a result, anti-human antibody-based methods are unsatisfactory. Most fail in reducing albumin and immunoglobulins, and some lead to a substantial removal of other highly abundant proteins, probably because of nonspecific interactions. A protein A/dye ligand-based method is efficient in reducing immunoglobulins, fibrinogen, and apolipoprotein A1 and A2, but not albumin, and protein identifications do not increase. Only peptide ligand libraries flatten the dynamic range, and increased protein identification (59.0%). Albumin and immunoglobulins are successfully depleted (60.7% and 35.9%, respectively). Although further studies will be required for reinforcing our observations, this work can provide a useful guide for cat plasma pretreatment in biomarker discovery studies.

DOI: https://doi.org/10.1002/pmic.201800191

Technical brief

All cats are grey in the dark: enrichment/depletion approaches for biomarker discovery on *Felis catus* plasma.

Running title: Enriching/depleting cat plasma samples

Laura Carcangiu[1], Salvatore Pisanu[2], Silvia Tore[3], Maria Filippa Addis[2,a], Eric Zini[1,4,5], Sergio Uzzau[2,b], Daniela Pagnozzi[2*]

1 Istituto Veterinario di Novara, Granozzo con Monticello, Novara, Italy
2 Porto Conte Ricerche, Science and Technology Park of Sardinia, Tramariglio, Alghero (Sassari), Italy
3 Sardegna Ricerche, Piscina Manna, Pula, Cagliari, Italy
4 Clinic for Small Animal Internal Medicine, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland
5 Department of Animal Medicine, Production and Health, University of Padova, Legnaro, Italy.
a Present address*: Department of Veterinary Medicine, University of Milan, Italy*
b Permanent address: *Department of Biomedical Sciences, University of Sassari, Sassari, Italy*

*Correspondence: Daniela Pagnozzi, PhD, Porto Conte Ricerche Srl, S.P. 55 Porto Conte - Capo Caccia, km 8,400 Loc. Tramariglio - 07041 Alghero (SS) – Italy. Tel: (+39) 079 998 400 – Fax: (+39) 079 998 567 (pagnozzi@portocontericerche.it).

Keywords: biomarker discovery; depletion; enrichment; plasma; shotgun proteomics.

Total number of words: 2406.

Abstract

In veterinary medicine, assay performance is often affected by the lack of species-specific diagnostic tools. Reliable biomarkers might be identified by investigating biological fluids of the species of interest, but protein sequence databases are often incomplete and human-specific devices for reducing sample complexity might fail when applied to animal plasma. Here, seven commercial methods, based on different capturing agents (anti-human antibodies, affinity ligands, mixture of antibodies and ligands, and combinatorial peptide ligand libraries), were applied to cat plasma and evaluated in terms of yield, identified proteins/ peptides, and relative abundance by high resolution shotgun proteomics and label-free quantitation. As a result, anti-human antibody-based methods were unsatisfactory. Most failed in reducing albumin and immunoglobulins, and some led to a substantial removal of other highly abundant proteins, probably because of nonspecific interactions. A protein A/dye ligand-based method was efficient in reducing immunoglobulins, fibrinogen and apolipoprotein A1 and A2, but not albumin, and protein identifications did not increase. Only peptide ligand libraries flattened the dynamic range, and increased protein identification (59.0%). Albumin and immunoglobulins were successfully depleted (60.7% and 35.9%, respectively). Although further studies will be required for reinforcing our observations, this work can provide a useful guide for cat plasma pretreatment in biomarker discovery studies.

Farm animal and pet welfare is the main goal of veterinary medicine for both economic and affection reasons. Nowadays, the performance of animal diagnostic serology is often sub-optimal, mostly due to the low availability of commercial species-specific immunoassays [1]. Even though clinical care may capitalize from what has already been discovered by human system biology, dedicated biomarker discovery and validation efforts would be needed for improving veterinary diagnostics [2]. Nevertheless, compared to humans, biomarker discovery studies directly performed on animal specimens lag considerably behind [3-5]. This might be partly due to specific difficulties in treating samples and to the limited availability of annotated protein sequence databases [6]. One of the main problems in proteomic analysis of biological specimens such as plasma, serum, and urine, is the presence of a huge dynamic molecular range of protein concentration [7-9]. The availability of commercial products aimed at depleting highly abundant protein and enriching the low abundant ones generally enables a more efficient and less challenging biomarker discovery process. However, most commercial solutions based on immunoaffinity criteria have been developed for human specimens, and only a limited amount of products are specific for, or compatible with some animal species [10-13]. Other products are composed by a mixture of affinity ligands (protein G or protein A, dye ligands, peptide libraries), theoretically allowing their use on a wider range of organisms [14-16]. To our knowledge, no antibody-based depletion kit has been declared by the manufacturers as suitable for feline samples, and only a limited number of papers report proteomic studies on cat biofluids [17], none of them applying procedures for depletion/enrichment of highly abundant proteins. Recently, a comparative evaluation of seven commercial products was performed by our group on a human serum sample by means of a shotgun proteomics approach [18, 19]. Here, we aimed at evaluating the same methods on a cat plasma sample to gather information on their depletion performances and to provide a selection guide for plasma biomarker discovery studies in this species.

Different capturing criteria were involved: i) immunoaffinity with antibodies directed against specific (human) proteins, ii) protein G and antibodies directed against human serum albumin (HSA), iii) protein A and cibacron blue dye-ligand, and iv) affinity for a combinatorial peptide ligand library. All products were applied to a pool of plasma samples and the enriched proteins were evaluated by SDS-PAGE, high resolution shotgun MS analysis and label-free quantitation. Plasma samples from four cats, were homogeneously pooled and stored at -80°C until use. Pool protein concentration was 63 mg/mL based on the BCA assay (Thermo Fisher Scientific, CA). Depletions and enrichments were performed according to the product manuals with the sample volumes recommended by the manufacturers [18]. Products were based on four different capturing agents: antibodies (Qproteome Albumin/IgG Depletion kit, "Qproteome", Qiagen, Sweden; ProteoPrep Immunoaffinity Albumin and IgG Depletion Kit, "ProteoPrep", Sigma-Aldrich, MO; Top 2 Abundant Protein Depletion Spin Columns "Top 2", and Top 12 Abundant Protein Depletion Spin Columns "Top 12", Thermo Fisher Scientific), specific ligands (Albumin/IgG Removal, "CibB-A", Thermo Fisher Scientific), mixture of antibodies and ligands (Albumin and IgG Depletion SpinTrap, "SpinTrap", GE Healthcare, Sweden), and combinatorial peptide ligand libraries (ProteoMiner beads, "ProteoMiner", Bio-Rad, CA), respectively. Table S1 summarizes the selected products and their relative features. Two technical replicates were performed for each procedure. A non-reducing Laemmli buffer was added to 10 µg of each depleted/enriched fraction and SDS-PAGE was carried out on any kD precast polyacrylamide gels using the Mini-Protean system (Bio-Rad). Filter aided sample preparation (FASP) was performed on all protein samples to obtain tryptic peptide mixtures; LC-MS/MS analysis was carried out on a Q Exactive mass spectrometer interfaced with an UltiMate 3000 RSLCnanoLC system (Thermo Fisher Scientific), by performing a linear gradient of 245 min from 5% to 37.5% of eluent B (0.1% formic acid in 80% acetonitrile) in eluent A (0.1% formic acid) at a flow rate of 250 nL/min. MS data were acquired using a data-dependent top12 method, as previously described [18, 21]. Proteome Discoverer

(version 1.4; Thermo Fisher Scientific) was used for protein identification with Sequest-HT as search engine. Raw files were processed with the following parameters: database *Felis silvestris catus,* downloaded from UniProtKB_TrEMBL (release 2017 08); enzyme trypsin, with two missed cleavages allowed; precursor mass tolerance 10 ppm; MS/MS tolerance 0.02 Da; charge states +2, +3, and +4; cysteine carbamidomethylation as static modification and methionine oxidation as dynamic modification. The percolator algorithm was used for protein significance and for peptide validation (FDR<0.01%). Uncharacterized proteins were identified by Basic Local Alignment Search Tool (BLAST) homology search.

Relative protein abundance was assessed by label-free quantitation based on the spectral counting (SpC) approach, calculating the normalized spectral abundance factor (NSAF) according to Old et al. [20]. All statistical analyses were performed using *t*-test (p-value ≤0.05).

According to the manufacturers suggestions, 10 µL of plasma were processed by Top 2, Top 12 and CibB-A columns, approximately corresponding to 0.63 mg of proteins; 25 µL of sample were incubated with SpinTrap, Qproteome and ProteoPrep columns, equivalent to 1.58 mg of proteins, and a 40 µL aliquot, accounting for 2.52 mg of proteins, was incubated with the appropriate volume of ProteoMiner beads, respecting the volume/volume ratio suggested by the manufacturer for serum/plasma samples. To evaluate protein yields for each product, the concentrations of depleted/enriched fraction were estimated. Unexpectedly, total recovery was very high in most of the applied protocols; more in detail, the protein content of recovered fractions spanned from a minimum of 31% for Top 12, to 33% for CibB-A, 35% for ProteoPrep, 39% for Top 2, up to a maximum of 48% for SpinTrap and Qproteome, when compared to the untreated plasma sample. Only ProteoMiner led to a very low yield of 1.8%, suggesting a depletion of a large amount of proteins. These yields were very different from the analogues calculated on human serum [18], revealing that both immuno- and chemical-affinity based products, except for the combinatorial peptide ligand library, have weaker interactions with cat plasma proteins.

To visualize the overall effect of each method, 10 µg aliquots of untreated plasma, of the depleted/enriched fractions and of the residual proteins were analyzed by SDS-PAGE. As shown in Figure 1, as a result of all procedures but ProteoMiner, the protein profile (red) remained very similar to the untreated sample (purple). However, a partial removal of highly abundant proteins did occur, as demonstrated by the residual protein fractions collected during each procedure (Figure 1, blue), which revealed a depletion of feline serum albumin (FSA, gel band at about 50 kDa) and immunoglobulins (Igs, gel bands from the top of the gel to about 100 kDa), even though in different amounts. After tryptic digestion, untreated plasma and depleted/enriched fractions were analyzed by a long-gradient LC-MS/MS to assess protein numbers and identities, and the relative abundance was estimated by label-free quantitation. Table 1 and Supplementary Table S2 summarize MS results. The highest number of proteins and peptides were identified after ProteoMiner enrichment, showing an increase of 59.0% and 24.9% when compared to untreated plasma, respectively. The other procedures led to a number of identified proteins quite similar to the undepleted plasma. Only CibB-A reduced the number of identified proteins and peptides, probably due to nonspecific interactions with many plasma proteins.

Further, to assess the effect of treatments on most abundant proteins, their variation in comparison to the untreated sample was calculated. Considering that the sequence identity between HSA and FSA is 82.4%, it would be reasonable to expect an efficient removal of this protein with an anti-HSA antibody, as reported by others for dog serum albumin (80.1% identity with HSA) [13]. Surprisingly, FSA removal was efficient only with the ProteoMiner approach, with a depletion of 60.7%; FSA content was slightly reduced by Top 2, whereas its amount was even increased by all the other products. CibB-A, despite exploiting a dye ligand bait supposed not to be species-specific, provided a 2.2-fold increase in the relative FSA content (Figure 2A). Regarding Igs, all the identifications referable to Igs were considered and the overall variation was calculated (Figure 2B). The best result was obtained by CibB-A, leading to a 59.7% of total Igs decrease, suggesting that

protein A has a higher affinity to cat Igs than protein G and anti-Igs antibodies. In general, all procedures allowed only a modest removal of Igs. Interestingly, results were similar to those seen with human serum [18] on other highly abundant proteins; more in detail, upon ProteoMiner treatment, serotransferrin, α2-macroglobulin, α1-antitrypsin, haptoglobin, and α1-acid glycoprotein (Accession: M3WBQ5, M3WMA9, M3WCX1, M3W4S3, and Q6KCA0, respectively) were decreased, whereas fibrinogen α-chain and fibrinogen β-chain (Accession: M3W022, and M3WII3) were increased; likewise, CibB-A led to a considerable reduction of fibrinogen. In addition, CibB-A induced a substantial decrease of apolipoprotein A1 and A2 (Accession: M3WPG6, and M3WN87) and an increase of serotransferrin, α2-macroglobulin, α1-antitrypsin, haptoglobin, and α1-acid glycoprotein contents. The other products, apart from few variations consisting in a slight decrease of fibrinogen α-chain and fibrinogen β-chain by Top 12, an increase of α1-antitrypsin by both Top 2 and Top 12, and of α1-acid glycoprotein by SpinTrap and QProteome, did not have an effect on the most abundant proteins. Further, the effect on low abundant proteins was evaluated; a total of 302 proteins were identified only by ProteoMiner approach (Supplementary Table S3). Among them, many human homologous proteins have been reported to be present in ng/mL (tenascin X, [22], thrombomodulin [23] and others) to pg/mL (inhibin B, [24]) concentrations in serum samples from normal subjects.

In conclusion, our data demonstrate that anti-human antibody-based methods are not suitable for cat plasma samples; on the contrary, CibB-A was very efficient in reducing Igs, fibrinogen and apolipoprotein A1 and A2; noteworthy, ProteoMiner allowed flattening of the sample dynamic range and a significant increase in the number of identified proteins. It is likely that the binding affinities of anti-human antibodies, protein G and cibacron blue dye-ligand are not adequate for cat proteins removal, even in case of high sequence identity with the homologous human protein, maybe due to a different protein domain exposure, or to the absence/presence of post-translational modifications that impairs the interactions. Otherwise, the high heterogeneity of ligands provided

by ProteoMiner beads (several millions of hexapeptide sequences) might favor the chance of establishinginteractions . Moreover, it should be underlined that all procedures, but ProteoMiner beads, provide for subsequent analysis a flow-through fraction, that may still be rich in target proteins due to their overabundance compared to the limited availability of binding sites; consequently, a fine optimization of the sample amount to be loaded should be performed before treatments. Although further studies carried out on a higher number of samples will be required to reinforce our observations, this work provides a useful guide for selecting depletion procedures in biomarker discovery studies on cat plasma.

This work was supported by Regione Autonoma della Sardegna, with the grants “Diagnostici per ANimali Da REddito e da Affezione”, D.AN.D.RE.A., and “art. 9 LR 20/2015”.
The authors declare no conflict of interest.
The authors thank Drs Tiziana Cubeddu and Giovanna Casalloni for providing plasma samples.

178 **References**

179 [1] T. Warnken, K Huber, K. Feige, BMC Vet. Res. 2016, 12, 196.

180 [2] R.E. Moore, J Kirwan, M.K. Doherty, P.D. Whitfield, Biomark. Insights 2007, 10 185.

181 [3] L. Galeandro, N.S. Sieber-Ruckstuhl, B. Riond, S. Hartnack, R. Hofmann-Lehmann, C.E.

182 Reusch, F.S. Boretti, J. Vet. Intern. Med. 2014, 28, 1433.

183 [4] R. Tanoue, I. Kume, Y. Yamamoto, K. Takaguchi, K. Nomiyama, S. Tanabe, T. Kunisue, J.

184 Chromatogr. A 2018, 2, 30.

185 [5] A. Tanca, D Pagnozzi, G.P. Burrai, M. Polinas, S Uzzau, E Antuofermo, M.F. Addis, J.

186 Proteomics 2012, 21, 561.

187 [6] F. Di Girolamo, A. D'Amato, I. Lante, F. Signore, M. Muraca, L. Putignani, Int. J. Mol. Sci.

188 2014, 15, 15396.

189 [7] G. Biosa, M.F. Addis, A. Tanca, S. Pisanu, T. Roggio, S. Uzzau, D. Pagnozzi, J. Proteomics

190 2011, 75, 93.

191 [8] A.H. Zhang, H. Sun, G.L. Yan, Y. Han, X.J. Wang, Appl. Biochem. Biotechnol. 2013, 170, 774.

192 [9] S. Filip, K. Vougas, J. Zoidakis, A. Latosinska, W. Mullen, G. Spasovski, H. Mischak, A.

193 Vlahou, J. Jankowski, PLoS One 2015, 10, e0133773.

194 [10] R. Günther, E. Krause, M. Schümann, J. Ausseil, J.M. Heard, I.E. Blasig, R.F. Haseloff, Fluids

195 Barriers CNS 2014, 23, 14.

196 [11] K. de Morais-Zani, K.F. Grego, A.S. Tanaka, A.M. Tanaka-Azevedo, J. Biomol. Tech. 2011,

197 22, 67.

198 [12] D.R. Haudenschild, A. Eldridge, P.J. Lein, B.A. Chromy, Biochem. Biophys. Res. Commun.

199 2014, 455, 84.

200 [13] M. Sundberg, J. Bergquist, M. Ramström, Biochem. Biophys. Rep. 2015, 3, 68.

201 [14] K. Björhall, T. Miliotis, P. Davidsson, Proteomics 2005, 5, 307.

[15] K. Leecharoenkiat, W. Sornjai, K. Khungwanmaythawee, A. Paemanee, C. Chaichana, S. Roytrakul, S. Fucharoen, S. Svasti, D.R. Smith, Dis. Markers 2014, 2014, 340214.

[16] A. D'Amato, A. Bachi, E. Fasoli, E. Boschetti, G. Peltre, H. Sénéchal, P.G. Righetti, J. Proteome Res. 2009, 8, 3925.

[17] M.D. Meachem, E.R. Snead, B.A. Kidney, M.L. Jackson, R. Dickinson, V. Larson, E. Simko, Can. J. Vet. Res. 2015, 79, 184.

[18] S. Pisanu, G. Biosa, L. Carcangiu, S. Uzzau, D. Pagnozzi, Talanta, 2018, 185, 213.

[19] S. Pisanu, G. Biosa, L. Carcangiu, S. Uzzau, D. Pagnozzi, Data Brief 2018, 19, 1765.

[20] W.M. Old, K. Meyer-Arendt, L. Aveline-Wolf, K.G. Pierce, A. Mendoza, J.R. Sevinsky, K.A. Resing, N.G. Ahn, Mol. Cell. Proteomics 2005, 4, 1487.

[21] D. Pagnozzi, F. Tamarozzi, A.M. Roggio, V. Tedde, M.F. Addis, S. Pisanu, G. Masu, C. Santucciu, A. Vola, A. Casulli, G. Masala, E. Brunetti, S. Uzzau, Clin. Infect. Dis. 2018, 66, 1342.

[22] K. Yamada, A. Watanabe, H. Takeshita, K. Matsumoto, Clin. Chim. Acta 2016, 459, 94.

[23] F. Califano, T. Giovanniello, P. Pantone, E. Campana, C. Parlapiano, F. Alegiani, G.M. Vincentelli, P. Turchetti, Eur. Rev. Med. Pharmacol. Sci. 2000, 4, 59.

[24] T.W. Kelsey, A. Miles, R.T. Mitchell, R.A. Anderson, W.H. Wallace, PLoS One 2016, 11, e0153843.

220 **Figure legends**

221 **Figure 1. SDS-PAGE of plasma fractions with and without treatments.** Protein profiles of
222 untreated plasma (purple), depleted/enriched fractions (red) and residual proteins (blue) after Top 2,
223 Top 12, SpinTrap, Qproteome, ProteoPrep, CibB-A, ProteoMiner. **U**: unbound protein fractions. **E**:
224 proteins eluted from the resins. **M**: Precision Plus molecular weight markers (Bio-Rad).

225

226 **Figure 2 Abundance (NSAF) variations of serum albumin (panel A) and total**
227 **immunoglobulins (panel B) after plasma treatment.** Plotted values represent the average
228 NSAF±SD. Asterisks indicate statistically significant differences between untreated and
229 depleted/enriched plasma according to t-test, with *p value≤ 0.05 and ** p value≤ 0.01.

230

231 **Table 1.** Numbers of identified proteins, peptides, PSMs and Search Inputs in untreated and
232 depleted/enriched plasma.
233

	Proteins	CV (%)	Peptides	CV (%)	PSMs	CV (%)	Search Inputs	CV (%)
Plasma	390 ± 8	1.99	4245 ± 63	1.48	13112 ± 86	0.65	80323 ± 86	0.11
CibB-A	264 ± 0**	0.00	2214 ± 30**	1.34	10455 ± 71**	0.68	70498 ± 1242**	1.76
ProteoPrep	353 ± 13	3.61	3932 ± 52*	1.33	12608 ± 66*	0.53	81329 ± 1114	1.37
SpinTrap	385 ± 16	4.04	4047 ± 65	1.61	13377 ± 213	1.60	78574 ± 1314	1.67
Top 12	395 ± 7	1.79	4303 ± 24	0.56	13690 ± 74*	0.54	79571 ± 1225	1.54
Qproteome	396 ± 6	1.61	3872 ± 69*	1.77	13488 ± 601	4.46	77057 ± 2502	3.25
Top 2	398 ± 7	1.78	4346 ± 13	0.29	13602 ± 71*	0.52	78296 ± 44**	0.06
ProteoMiner	620 ± 2**	0.34	5303 ± 42**	0.80	15408 ± 65**	0.42	77151 ± 619*	0.80

234 Numbers represent the average from technical replicates ± SD.
235 CV: coefficient of variation. Significant differences between untreated and depleted/enriched plasma, according to *t*-test, are
236 indicated as * (p value≤0.05), or ** (p value≤0.01).
237